

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/615,285DATE: 07/25/2000
TIME: 14:41:55Input Set : A:\seqlist.txt
Output Set: N:\CRF3\07252000\I615285.raw

4 <110> APPLICANT: Daniel E.H. Afar
5 Rene S. Hubert
6 Kahan Leong
7 Arthur B. Raitano
8 Douglas C. Saffran
9 Stephen C. Mitchell
10 Aya Jakobovits
11 Mary Faris
12 Igor Vivanco
14 <120> TITLE OF INVENTION: NOVEL TUMOR ANTIGEN USEFUL IN DIAGNOSIS
15 AND THERAPY OF PROSTATE AND COLON CANCER
18 <130> FILE REFERENCE: 129.9US11
C--> 20 <140> CURRENT APPLICATION NUMBER: US/09/615,285
C--> 20 <141> CURRENT FILING DATE: 2000-07-12
20 <150> PRIOR APPLICATION NUMBER: 09/323,597
21 <151> PRIOR FILING DATE: 1999-06-01
23 <160> NUMBER OF SEQ ID NOS: 29
25 <170> SOFTWARE: FastSEQ for Windows Version 4.0
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 1738
29 <212> TYPE: DNA
30 <213> ORGANISM: Homo Sapiens
32 <220> FEATURE:
33 <221> NAME/KEY: CDS
34 <222> LOCATION: (112)...(1588)
36 <400> SEQUENCE: 1
37 ggcggaggcg gagcgagg ggcggggcg gggagcgccg cctggagcgc ggcaggtcat 60
38 attgaacatt ccagatacct atcattactc gatgctgttg ataacagcaa g atg gct 117
39 Met Ala
40 1
42 ttg aac tca ggg tca cca cca gct att gga cct tac tat gaa aac cat 165
43 Leu Asn Ser Gly Ser Pro Pro Ala Ile Gly Pro Tyr Tyr Glu Asn His
44 5 10 15
46 gga tac caa ccg gaa aac ccc tat ccc gca cag ccc act gtg gtc ccc 213
47 Gly Tyr Gln Pro Glu Asn Pro Tyr Pro Ala Gln Pro Thr Val Val Pro
48 20 25 30
50 act gtc tac gag gtg cat ccg gct cag tac tac ccg tcc ccc gtg ccc 261
51 Thr Val Tyr Glu Val His Pro Ala Gln Tyr Tyr Pro Ser Pro Val Pro
52 35 40 45 50
54 cag tac gcc ccg agg gtc ctg acg cag gct tcc aac ccc gtc gtc tgc 309
55 Gln Tyr Ala Pro Arg Val Leu Thr Gln Ala Ser Asn Pro Val Val Cys
56 55 60 65
58 acg cag ccc aaa tcc cca tcc ggg aca gtg tgc acc tca aag act aag 357
59 Thr Gln Pro Lys Ser Pro Ser Gly Thr Val Cys Thr Ser Lys Thr Lys
60 70 75 80
62 aaa gca ctg tgc atc acc ttg acc ctg ggg acc ttc ctc gtg gga gct 405
63 Lys Ala Leu Cys Ile Thr Leu Thr Leu Gly Thr Phe Leu Val Gly Ala

ENTERED

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/615,285

DATE: 07/25/2000
 TIME: 14:41:55

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\07252000\I615285.raw

64	85	90	95	
66	gcg ctg gcc gct ggc cta ctc tgg aag ttc atg ggc agc aag tgc tcc	453		
67	Ala Leu Ala Ala Gly Leu Leu Trp Lys Phe Met Gly Ser Lys Cys Ser			
68	100 105 110			
70	aac tct ggg ata gag tgc gac tcc tca ggt acc tgc atc aac ccc tct	501		
71	Asn Ser Gly Ile Glu Cys Asp Ser Ser Gly Thr Cys Ile Asn Pro Ser			
72	115 120 125 130			
74	aac tgg tgt gat ggc gtg tca cac tgc ccc ggc ggg gag gac gag aat	549		
75	Asn Trp Cys Asp Gly Val Ser His Cys Pro Gly Gly Glu Asp Glu Asn			
76	135 140 145			
78	cgg tgt gtt cgc ctc tac gga cca aac ttc atc ctt cag gtg tac tca	597		
79	Arg Cys Val Arg Leu Tyr Gly Pro Asn Phe Ile Leu Gln Val Tyr Ser			
80	150 155 160			
82	tct cag agg aag tcc tgg cac cct gtg tgc caa gac gac tgg aac gag	645		
83	Ser Gln Arg Lys Ser Trp His Pro Val Cys Gln Asp Asp Trp Asn Glu			
84	165 170 175			
86	aac tac ggg cgg gcg gcc tgc agg gac atg ggc tat aag aat aat ttt	693		
87	Asn Tyr Gly Arg Ala Ala Cys Arg Asp Met Gly Tyr Lys Asn Asn Phe			
88	180 185 190			
90	tac tct agc caa gga ata gtg gat gac agc gga tcc acc agc ttt atg	741		
91	Tyr Ser Ser Ser Gln Gly Ile Val Asp Asp Ser Gly Ser Thr Ser Phe Met			
92	195 200 205 210			
94	aaa ctg aac aca agt gcc ggc aat gtc gat atc tat aaa aaa ctg tac	789		
95	Lys Leu Asn Thr Ser Ala Gly Asn Val Asp Ile Tyr Lys Lys Leu Tyr			
96	215 220 225			
98	cac agt gat gcc tgt tct tca aaa gca gtg gtt tct tta cgc tgt ata	837		
99	His Ser Asp Ala Cys Ser Ser Lys Ala Val Val Ser Leu Arg Cys Ile			
100	230 235 240			
102	gcc tgc ggg gtc aac ttg aac tca agc cgc cag agc agg att gtg ggc	885		
103	Ala Cys Gly Val Asn Leu Asn Ser Ser Arg Gln Ser Arg Ile Val Gly			
104	245 250 255			
106	ggc gag agc gcg ctc ccg ggg gcc tgg ccc tgg cag gtc agc ctg cac	933		
107	Gly Glu Ser Ala Leu Pro Gly Ala Trp Pro Trp Gln Val Ser Leu His			
108	260 265 270			
110	gtc cag aac gtc cac gtg tgc gga ggc tcc atc atc acc ccc gag tgg	981		
111	Val Gln Asn Val His Val Cys Gly Gly Ser Ile Ile Thr Pro Glu Trp			
112	275 280 285 290			
114	atc gtg aca gcc gcc cac tgc gtg gaa aaa cct ctt aac aat cca tgg	1029		
115	Ile Val Thr Ala Ala His Cys Val Glu Lys Pro Leu Asn Asn Pro Trp			
116	295 300 305			
118	cat tgg acg gca ttt gcg ggg att ttg aga caa tct ttc atg ttc tat	1077		
119	His Trp Thr Ala Phe Ala Gly Ile Leu Arg Gln Ser Phe Met Phe Tyr			
120	310 315 320			
122	gga gcc gga tac caa gta gaa aaa gtg att tct cat cca aat tat gac	1125		
123	Gly Ala Gly Tyr Gln Val Glu Lys Val Ile Ser His Pro Asn Tyr Asp			
124	325 330 335			
126	tcc aag acc aag aac aat gac att gcg ctg atg aag ctg cag aag cct	1173		
127	Ser Lys Thr Lys Asn Asn Asp Ile Ala Leu Met Lys Leu Gln Lys Pro			
128	340 345 350			

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/615,285
 DATE: 07/25/2000
 TIME: 14:41:55

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\07252000\I615285.raw

```

130 ctg act ttc aac gac cta gtg aaa cca gtg tgt ctg ccc aac cca ggc 1221
131 Leu Thr Phe Asn Asp Leu Val Lys Pro Val Cys Leu Pro Asn Pro Gly
132 355 360 365 370
134 atg atg ctg cag cca gaa cag ctc tgc tgg att tcc ggg tgg ggg gcc 1269
135 Met Met Leu Gln Pro Glu Gln Leu Cys Trp Ile Ser Gly Trp Gly Ala
136 375 380 385
138 acc gag gag aaa ggg aag acc tca gaa gtg ctg aac gct gcc aag gtg 1317
139 Thr Glu Glu Lys Gly Lys Thr Ser Glu Val Leu Asn Ala Ala Lys Val
140 390 395 400
142 ctt ctc att gag aca cag aga tgc aac agc aga tat gtc tat gac aac 1365
143 Leu Leu Ile Glu Thr Gln Arg Cys Asn Ser Arg Tyr Val Tyr Asp Asn
144 405 410 415
146 ctg atc aca cca gcc atg atc tgt gcc gcc ttc ctg cag ggg aac gtc 1413
147 Leu Ile Thr Pro Ala Met Ile Cys Ala Gly Phe Leu Gln Gly Asn Val
148 420 425 430
150 gat tct tgc cag ggt gac agt gga ggg cct ctg gtc act tgc aag aac 1461
151 Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Ser Lys Asn
152 435 440 445 450
154 aat atc tgg tgg ctg ata ggg gat aca agc tgg ggt tct ggc tgt gcc 1509
155 Asn Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly Cys Ala
156 455 460 465
158 aaa gct tac aga cca gga gtg tac ggg aat gtg atg gta ttc acg gac 1557
159 Lys Ala Tyr Arg Pro Gly Val Tyr Gly Asn Val Met Val Phe Thr Asp
160 470 475 480
162 tgg att tat cga caa atg agg gca gac gcc t aatccacatg gtcttcgtcc 1608
W--> 163 Trp Ile Tyr Arg Gln Met Arg Ala Asp Gly
164 485 490
166 ttgacgtcgt tttaacaagaa aacaatgggg ctggttttgc ttccccgtgc atgatttact 1668
167 cttagagatg attcagaggt cacttcattt ttattaaaca gtgaacttgt ctggcaaaaa 1728
168 aaaaaaaaaa 1738
170 <210> SEQ ID NO: 2
171 <211> LENGTH: 492
172 <212> TYPE: PRT
173 <213> ORGANISM: Homo Sapiens
175 <400> SEQUENCE: 2
176 Met Ala Leu Asn Ser Gly Ser Pro Pro Ala Ile Gly Pro Tyr Tyr Glu
177 1 5 10 15
178 Asn His Gly Tyr Gln Pro Glu Asn Pro Tyr Pro Ala Gln Pro Thr Val
179 20 25 30
180 Val Pro Thr Val Tyr Glu Val His Pro Ala Gln Tyr Tyr Pro Ser Pro
181 35 40 45
182 Val Pro Gln Tyr Ala Pro Arg Val Leu Thr Gln Ala Ser Asn Pro Val
183 50 55 60
184 Val Cys Thr Gln Pro Lys Ser Pro Ser Gly Thr Val Cys Thr Ser Lys
185 65 70 75 80
186 Thr Lys Lys Ala Leu Cys Ile Thr Leu Thr Leu Gly Thr Phe Leu Val
187 85 90 95
188 Gly Ala Ala Leu Ala Ala Gly Leu Leu Trp Lys Phe Met Gly Ser Lys
189 100 105 110

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/615,285
 DATE: 07/25/2000
 TIME: 14:41:55

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\07252000\I615285.raw

```

190 Cys Ser Asn Ser Gly Ile Glu Cys Asp Ser Ser Gly Thr Cys Ile Asn
191      115      120      125
192 Pro Ser Asn Trp Cys Asp Gly Val Ser His Cys Pro Gly Gly Glu Asp
193      130      135      140
194 Glu Asn Arg Cys Val Arg Leu Tyr Gly Pro Asn Phe Ile Leu Gln Val
195      145      150      155      160
196 Tyr Ser Ser Gln Arg Lys Ser Trp His Pro Val Cys Gln Asp Asp Trp
197      165      170      175
198 Asn Glu Asn Tyr Gly Arg Ala Ala Cys Arg Asp Met Gly Tyr Lys Asn
199      180      185      190
200 Asn Phe Tyr Ser Ser Gln Gly Ile Val Asp Asp Ser Gly Ser Thr Ser
201      195      200      205
202 Phe Met Lys Leu Asn Thr Ser Ala Gly Asn Val Asp Ile Tyr Lys Lys
203      210      215      220
204 Leu Tyr His Ser Asp Ala Cys Ser Ser Lys Ala Val Ser Leu Arg
205      225      230      235      240
206 Cys Ile Ala Cys Gly Val Asn Leu Asn Ser Ser Arg Gln Ser Arg Ile
207      245      250      255
208 Val Gly Gly Glu Ser Ala Leu Pro Gly Ala Trp Pro Trp Gln Val Ser
209      260      265      270
210 Leu His Val Gln Asn Val His Val Cys Gly Gly Ser Ile Ile Thr Pro
211      275      280      285
212 Glu Trp Ile Val Thr Ala Ala His Cys Val Glu Lys Pro Leu Asn Asn
213      290      295      300
214 Pro Trp His Trp Thr Ala Phe Ala Gly Ile Leu Arg Gln Ser Phe Met
215      305      310      315      320
216 Phe Tyr Gly Ala Gly Tyr Gln Val Glu Lys Val Ile Ser His Pro Asn
217      325      330      335
218 Tyr Asp Ser Lys Thr Lys Asn Asn Asp Ile Ala Leu Met Lys Leu Gln
219      340      345      350
220 Lys Pro Leu Thr Phe Asn Asp Leu Val Lys Pro Val Cys Leu Pro Asn
221      355      360      365
222 Pro Gly Met Met Leu Gln Pro Glu Gln Leu Cys Trp Ile Ser Gly Trp
223      370      375      380
224 Gly Ala Thr Glu Glu Lys Gly Lys Thr Ser Glu Val Leu Asn Ala Ala
225      385      390      395      400
226 Lys Val Leu Leu Ile Glu Thr Gln Arg Cys Asn Ser Arg Tyr Val Tyr
227      405      410      415
228 Asp Asn Leu Ile Thr Pro Ala Met Ile Cys Ala Gly Phe Leu Gln Gly
229      420      425      430
230 Asn Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Ser
231      435      440      445
232 Lys Asn Asn Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly
233      450      455      460
234 Cys Ala Lys Ala Tyr Arg Pro Gly Val Tyr Gly Asn Val Met Val Phe
235      465      470      475      480
236 Thr Asp Trp Ile Tyr Arg Gln Met Arg Ala Asp Gly
237      485      490
239 <210> SEQ ID NO: 3

```

RAW SEQUENCE LISTING DATE: 07/25/2000
 PATENT APPLICATION: US/09/615,285 TIME: 14:41:55

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\07252000\I615285.raw

```

240 <211> LENGTH: 2479
241 <212> TYPE: DNA
242 <213> ORGANISM: Homo sapiens
244 <220> FEATURE:
245 <221> NAME/KEY: CDS
246 <222> LOCATION: (57)...(1534)
248 <400> SEQUENCE: 3
249 gtcattattga acattccaga tacctatcat tactcgatgc tgttgataac agcaag atg      59
250                                     Met
251                                     1
253 gct ttg aac tca ggg tca cca cca gct att gga cct tac tat gaa aac      107
254 Ala Leu Asn Ser Gly Ser Pro Pro Ala Ile Gly Pro Tyr Tyr Glu Asn
255                                     5      10      15
257 cat gga tac caa ccg gaa aac ccc tat ccc gca cag ccc act gtg gtc      155
258 His Gly Tyr Gln Pro Glu Asn Pro Tyr Pro Ala Gln Pro Thr Val Val
259                                     20      25      30
261 ccc act gtc tac gag gtg cat ccg gct cag tac tac ccg tcc ccc gtg      203
262 Pro Thr Val Tyr Glu Val His Pro Ala Gln Tyr Tyr Pro Ser Pro Val
263                                     35      40      45
265 ccc cag tac gcc ccg agg gtc ctg acg cag gct tcc aac ccc gtc gtc      251
266 Pro Gln Tyr Ala Pro Arg Val Leu Thr Gln Ala Ser Asn Pro Val Val
267 50                                     55      60      65
269 tgc acg cag ccc aaa tcc cca tcc ggg aca gtg tgc acc tca aag act      299
270 Cys Thr Gln Pro Lys Ser Pro Ser Gly Thr Val Cys Thr Ser Lys Thr
271                                     70      75      80
273 aag aaa gca ctg tgc atc acc ttg acc ctg ggg acc ttc ctc gtg gga      347
274 Lys Lys Ala Leu Cys Ile Thr Leu Thr Leu Gly Thr Phe Leu Val Gly
275                                     85      90      95
277 gct gcg ctg gcc gct ggc cta ctc tgg aag ttc atg ggc agc aag tgc      395
278 Ala Ala Leu Ala Ala Gly Leu Leu Trp Lys Phe Met Gly Ser Lys Cys
279 100      105      110
281 tcc aac tct ggg ata gag tgc gac tcc tca ggt acc tgc atc aac ccc      443
282 Ser Asn Ser Gly Ile Glu Cys Asp Ser Ser Gly Thr Cys Ile Asn Pro
283 115      120      125
285 tct aac tgg tgt gat ggc gtg tca cac tgc ccc ggc ggg gag gac gag      491
286 Ser Asn Trp Cys Asp Gly Val Ser His Cys Pro Gly Gly Glu Asp Glu
287 130      135      140      145
289 aat cgg tgt gtt cgc ctc tac gga cca aac ttc atc ctt cag atg tac      539
290 Asn Arg Cys Val Arg Leu Tyr Gly Pro Asn Phe Ile Leu Gln Met Tyr
291 150      155      160
293 tca tct cag agg aag tcc tgg cac cct gtg tgc caa gac gac tgg aac      587
294 Ser Ser Gln Arg Lys Ser Trp His Pro Val Cys Gln Asp Asp Trp Asn
295 165      170      175
297 gag aac tac ggg cgg gcg gcc tgc agg gac atg ggc tat aag aat aat      635
298 Glu Asn Tyr Gly Arg Ala Ala Cys Arg Asp Met Gly Tyr Lys Asn Asn
299 180      185      190
301 ttt tac tct agc caa gga ata gtg gat gac agc gga tcc acc agc ttt      683
302 Phe Tyr Ser Ser Gln Gly Ile Val Asp Asp Ser Gly Ser Thr Ser Phe
303 195      200      205

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY DATE: 07/25/2000
PATENT APPLICATION: US/09/615,285 TIME: 14:41:56

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\07252000\I615285.raw

L:20 M:270 C: Current Application Number differs, Replaced Current Application No
L:20 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:163 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 1
L:374 M:361 W: Invalid Split Codon, Sequence data for SEQ ID#: 3
L:472 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:5
L:472 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:5
L:472 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:5
L:472 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:5
L:472 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:5
L:473 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:5
L:473 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:5
L:473 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:5
L:473 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:5
M:340 Repeated in SeqNo=5